

高速 LED 背光速视器揭示拓扑性质的快速加工 *

钱晨灿^{1, 2)} 刘祖祥^{1)**}

(¹⁾ 中国科学院生物物理研究所, 脑与认知科学国家重点实验室, 北京 100101; ²⁾ 中国科学院大学, 北京 100049)

摘要 灵长类视觉系统能够非常有效地从环境中提取稳定的不变性特征, 而不受各种变换与干扰的影响. 这种不变性知觉的时间与动态过程为视觉理论提供了重要的约束. 我们开发了一种高速发光二极管(LED)背光速视器, 能够短暂呈现 1 ms 的视觉刺激, 并以亚毫秒级精度调整刺激呈现时间. 这使得我们可以通过控制视觉输入的可获取时间, 估计感觉线索积累的心理物理曲线, 从而研究在图形结构优势效应任务中, 拓扑、射影、仿射和欧氏不变性的相对加工速度. 实验结果表明拓扑不变性所需要的呈现时间最短, 这与大范围首先理论的预测相一致.

关键词 图形结构优势效应, 拓扑不变性, 时间过程, 心理物理曲线, 速视器

学科分类号 B842.1

DOI: 10.16476/j.pibb.2015.0300

灵长类视觉系统具有极为出色的空间形状分析处理能力^[1], 可以快速在复杂场景中特异性地识别物体, 而不受其位置、大小和视角等变化的影响^[2]. 这种不变性知觉能力被认为主要依赖大脑腹侧视觉通路的层次化加工^[3]. 腹侧通路脑区按照投射关系的层次由低到高, 感受野逐渐增大, 选择的特征趋于复杂而不变性增强^[4-5]. 因此一种观点认为不变性知觉可通过从局部开始, 以前馈为主的表征变换来实现^[4, 6]. 近期基于深度神经网络的人工视觉系统获得了很大进步, 其数据和理论也支持这种观点^[7-8]. 然而大量心理学实验表明, 在知觉决策任务中, 人对物体和场景的整体性质反应更快, 所需整合时间更短, 引起的注意、学习等效应更强^[9-10]. 一个典型示例是图形结构优势效应(configural superiority effect, CSE)^[11-12]: 加入不包含额外位置信息的图形与原有图形所构成的整体, 比原有图形的搜索反应时更短. 有两类理论试图调和这种看似矛盾的结果. “优先使用”理论认为, 整体性质同样需要从局部开始逐渐加工, 但它们更早地被后续认知过程所访问和使用, 导致了更快地反应. Pomerantz 的密封通道理论(sealed channel hypothesis)^[13], 以及 Hochstein 和 Ahissar 的反向层次理论(reverse hierarchy theory)^[9]属于这一类. “优先加工”理论认为, 某些整体性质可以经由快速通

道优先得加工, 而不依赖于局部性质加工的结果. Bar 等^[14]提出空间低频信息可经由背侧视觉通路或皮层下通路优先加工, 并通过眶额叶皮层(orbital frontal cortex, OFC)反馈到腹侧通路指导物体识别. 陈霖的大范围首先(global first)理论^[10]认为: 图形的各种性质在依照 Erlangen 纲领呈依次包含关系的欧氏、仿射、射影和拓扑等一系列几何变换下具有不同的稳定性; 越稳定的性质对于生物体而言可能具有更重要的意义, 从而需要更快地加工和使用(称为功能层次); 其中最稳定的二维拓扑性质可能经由皮层下通路优先加工, 并在前颞叶(anterior temporal lobe, ATL)形成由拓扑不变性定义的知觉物体表征, 引导后续的各种认知过程^[15]. 研究不同性质加工所需要的通路、相互依赖关系和完成的时间是区分上述理论的关键.

反应时是研究认知过程动态特性的重要心理学指标, 特别是对于远高于阈值或者受试极少犯错的任务^[16-17]. 对于 CSE, 大范围首先理论的预言

* 国家重点基础研究发展计划(973)(2012CB825500, 2015CB351701), 国家自然科学基金(91132302)和中国科学院基金(XDB02010001, XDB02050001)资助项目.

** 通讯联系人.

Tel: 010-64888565-808, E-mail: zxliu@bcslab.ibp.ac.cn

收稿日期: 2015-09-29, 接受日期: 2016-01-07

是：不同稳定性的不变性具有不同的结构优势，而以拓扑性质差异的优势最大，反应时最短，这与实验观察到的结果一致^[18]。然而，反应时实验具有两个重要的缺点，使其难以真正比较不同性质的加工所需要的时间。第一个缺点来自反应时研究中的可比性假设^[17]，即不同实验条件只在我们感兴趣的认知过程上不同，而同样构成反应时一部分的其他认知过程完全相同。在 CSE 任务里，受试需要指出哪一个象限包含的图形与其他象限不同，图形的异同在不同的实验条件下由不同的几何不变性定义。这里我们感兴趣的是几何不变性的抽提过程，然而为了做出反应，受试还需要定位出那个不同的象限。由于 4 个象限间图形的组织效应(grouping effect)强弱会影响定位过程的速度^[19]，而不同条件下组织效应可能会有差异，这就导致反应时的差别可能反映了几何不变性抽提时间的不同，也可能反映的是象限间组织效应的强弱。第二个缺点来自速度-准确度权衡(speed-accuracy tradeoff, SAT)^[20-21]现象，即受试可以在完成任务的速度和准确度之间以牺牲一方为代价换取另一方的提升，这种交换可以是有意识的，也可能是无意识的。这意味着本质上需要同时测量反应时和正确率这两个指标(即 SAT 曲线上的一个点)才能确定受试的任务表现或者能力，而整条 SAT 曲线代表相同的能力但具有不同的偏好，类似于信号检测理论中 ROC 曲线对敏感性和偏好的刻画。考虑到受试对速度和准确度的偏好可能随实验条件而改变(无论受试本人是否意识到这一点)，我们无法单纯根据反应时的差异推断不同几何不变性加工时间的长度，除非正确率也表现出“相同方向”的差异，例如，反应时短的正确率也高。即使在这种情况下，由于反应时实验中每个条件的试次数通常相对较少，不容易得到对正确率的准确估计。同时反应时实验中任务的正确率往往很高(例如，95%以上)，位于 SAT 曲线正确率趋于饱和的一端，这意味着受试内在偏好的改变可以在几乎不改变正确率的情况下极大地影响反应时。综上所述，由于这两个缺点的存在，利用反应时结果推断几何不变性的加工时间需要十分小心。

Stanford 等^[22]在猕猴的眼动反应上成功地运用了一种称为强迫反应范式(compelled-response paradigm)的新方法，一定程度上克服了反应时实验的缺点。受试被要求在看到一个额外的启动信号后立即做出反应，无论此时真正需要判断的实验刺激是否已经呈现。根据实验刺激距离启动信号时间

差的不同(可正可负)，受试的正确率可以从随机水平升高到接近饱和。根据这条时间响应曲线(tachometric curve)，可以较为精确地推断实验刺激的加工速度。然而强迫反应范式也有两个缺点阻碍了其广泛应用。首先，为了较为准确地估计时间响应曲线，需要大量的试次。在 Stanford 等报道的实验中，得到一条曲线每只猕猴大约需要完成 5 000~6 000 个试次，这在人类受试上较难实现。其次，受试如果在“随机猜测”和“等待刺激”两种策略中随机摇摆，而不严格遵守实验的指导语，最后平均的结果也会产生类似的时间响应曲线，这给实验结果的解释带来了一定的困难。

研究认知过程动态的第三种方法是通过限制刺激的加工时间，例如短暂呈现或者在刺激前后加上掩蔽刺激^[23]，来绘制心理物理曲线。这种心理物理方法也能有效地克服反应时实验的缺点。首先，由于我们感兴趣的感知觉信息的可获取时间是由实验操作系统性控制的，其他无关认知过程所需时间在条件间的差异不会影响结果。其次，由于测量的是正确率或者敏感性(d')，对反应速度没有要求，受试同样处于 SAT 曲线的正确率饱和区域，其内在偏好的变化对正确率的影响很小，也就不会影响实验结果。然而，这种心理物理方法也有一个局限性，就是对知觉可分辨性高的刺激时间分辨力有限。加工时间稍微改变一点，正确率就迅速变化，在实际中难以操作。在 CSE 实验中所使用的刺激，特别是包含拓扑性质差异的条件，就属于这种情况，使得实验在传统视觉呈现设备上难以开展。

为此，我们设计并实现了一种基于发光二极管(LED)背光的新型速视器，能够以 1 ms 甚至更小的时间间隔精确地改变视觉刺激呈现的时间，最短呈现时间也可达到 1 ms 甚至更短，从而可以精细地绘制高可分辨性刺激的心理物理曲线。我们利用这套新型速视器和心理物理方法研究了在 CSE 任务中，欧氏、仿射、射影以及拓扑不变性加工速度的差异。结果表明，拓扑性质的处理速度最快。

1 LED 背光速视器

1.1 背景与相关工作

传统的阴极射线管(CRT)显示器刷新率一般为 100 Hz，最高可达 200 Hz，这意味着刺激呈现时间只能以 10 ms 或者 5 ms 为单位进行调整，难以满足本实验的需求。此外，CRT 的扫描工作模式使得屏幕上不同位置的刺激呈现时间不完全同步，左

上角和右下角刺激呈现的时间最长可能相差 10 ms. 这对于本实验在 4 个象限同时呈现刺激, 并精确控制呈现时间尤为不利.

液晶显示器(LCD)技术的性能在不断提升, 目前 3D LCD 的刷新率已经可以达到 144 Hz, 灰度值转换的响应时间也可达到 1 ms, 但要实现亚毫秒级精度的呈现时间控制尚有难度. 同时, 虽然显示在 LCD 上的静态图像达到稳态后不需要扫描, 但是在刺激刚刚更新时, 仍然存在一个串行刷新的过程, 而非同步更新.

最新的有机发光二极管(OLED)显示器直接采用 LED 作为显示元件, 具有高速的开关响应特性和极高的对比度. 然而, 目前主流的 OLED 显示器仅有 60 Hz 的刷新率, 无法满足实验需求.

传统的机械电子速视器结合机械快门和闪光灯, 能实现最短约 5 ms 的呈现时间, 依然难以满足本实验的需求. 而且其刺激图形和呈现时间的改变都需要手动进行(更换卡片、调整旋钮), 相比现在计算机化的刺激呈现设备, 实验过程十分不便.

Thurgood 等^[24]开发了基于 LED 背光的新型速视器. 他们采用 LCD 面板作为图形显示设备, 并用自制的 LED 阵列作为背光源, LED 的开关由单片机控制. 该方案的核心在于, LCD 上图形的可见性取决于 LED 背光是否打开, 所以尽管 LCD 只能以 16.7 ms 的整数倍时长呈现刺激, 且图形切换后可能需要经过多达 50 ms 灰度值才真正稳定, 但由于 LED 的开关响应速度和单片机的时间控制精度都可以达到亚毫秒级, 只要在 LCD 面板图形更新达到稳态后, 再短暂地打开 LED 背光, 就可以精确实现视觉刺激短至 1 ms 的同步呈现. 然而由于他们的设计中只包含一块屏幕, 而在刺激短暂呈现的前后都必须关闭 LED 背光, 因此视觉刺激必然呈现为黑暗环境下的一次闪光, 缺少时间上前后相邻的掩蔽刺激. 这导致物理刺激虽然很快消失, 但视觉短时记忆可以不受干扰地维持较长的时间, 事实上刺激的加工时间并未得到有效的控制. 在他们的动物图片识别命名实验中, 呈现时间即使短至 1 ms, 正确率也高达 83%^[25], 而类似的任务如果采用动态掩蔽刺激, 所需的最短呈现时间大约在 20 ms 以上^[23].

Sperdin 等^[26-27]采用了基于两块 LCD 面板和一面半反半透镜的设计, 同样使用单片机独立控制两块面板的 LED 背光作为开关. 在通常的工作模式

下, 一块面板的背光打开, 另一块面板的背光就同时关闭, 两块屏幕切换显示. 在一块屏幕背光短暂打开呈现刺激的前后, 受试看到的不是黑屏, 而是另一块屏幕上的内容, 这就让呈现时间上相邻的静态甚至动态掩蔽刺激成为可能. 我们的方案与 Sperdin 等人类似, 但更为简单.

1.2 结构与原理

我们的速视器系统由一块 LCD 面板, 两块 LED 背光和一面半反半透镜组成, 如图 1 所示. LED1 和 LCD 组成屏幕 1 用于短暂呈现需要在不同试次中变化的实验刺激, LED2 和一块用于放置投影胶片的玻璃板组成屏幕 2 用于呈现在实验中无需更换的刺激(如注视点). 两块屏幕垂直位于速视器的背部与顶部, 半反半透镜以 45°角倾斜置于两者之间, 使得从观察窗口看上去只有一块视距约为 75 cm 的屏幕. LCD 面板尺寸为 41 cm×23 cm, 分辨率为 1366×768. 在 2 块屏幕表面都加装了方形光阑, 通过改变其大小和位置可以调整 2 块屏幕边缘的对齐. 光阑尺寸约为 28 cm×17 cm, 所以速视器的有效视角约为 21°×13°.

屏幕 1 由飞利浦的 191EL2 商用液晶显示器改装, 标称刷新率是 60 Hz, 响应时间为 5 ms. 我们确定了其内部用于给 LED1 供电的电源线路, 采用 IRF540 低导通电阻功率型金属氧化物半导体场效应管(MOSFET)对其进行高速开关控制. 由于 LED 驱动电路内部具有反馈保护功能, 驱动电流的突然中断会导致振荡升压电路停止工作, 因此使用一组大功率电阻作为匹配负载, 在 LED1 关闭期间将电流切换至负载电阻, 保持驱动电路的稳定工作. LED2 是一个自制的 LED 阵列, 其驱动电流大小由三极管控制, 可通过旋转电位器进行调节, 进而调节背光亮亮度.

LCD 面板通过标准 VGA 线与计算机显卡相连, 始终正常显示显卡输出的内容, 但是屏幕内容的可见性由 LED 背光控制. 两块 LED 背光的开关反相耦合, 即 LED1 打开, 则同时 LED2 关闭. 当 LED1 打开时, LCD 面板上显示的内容透过半反半透镜可见; 当 LED2 打开时, 投影胶片上的内容经半反半透镜反射可见. 背光开关由 Arduino Uno 单片机控制, 单片机通过 USB 线接受计算机串口通信指令的控制. 背光状态翻转时, 单片机同时向外发送触发信号, 可与脑电图、眼动或者反应时记录设备同步.

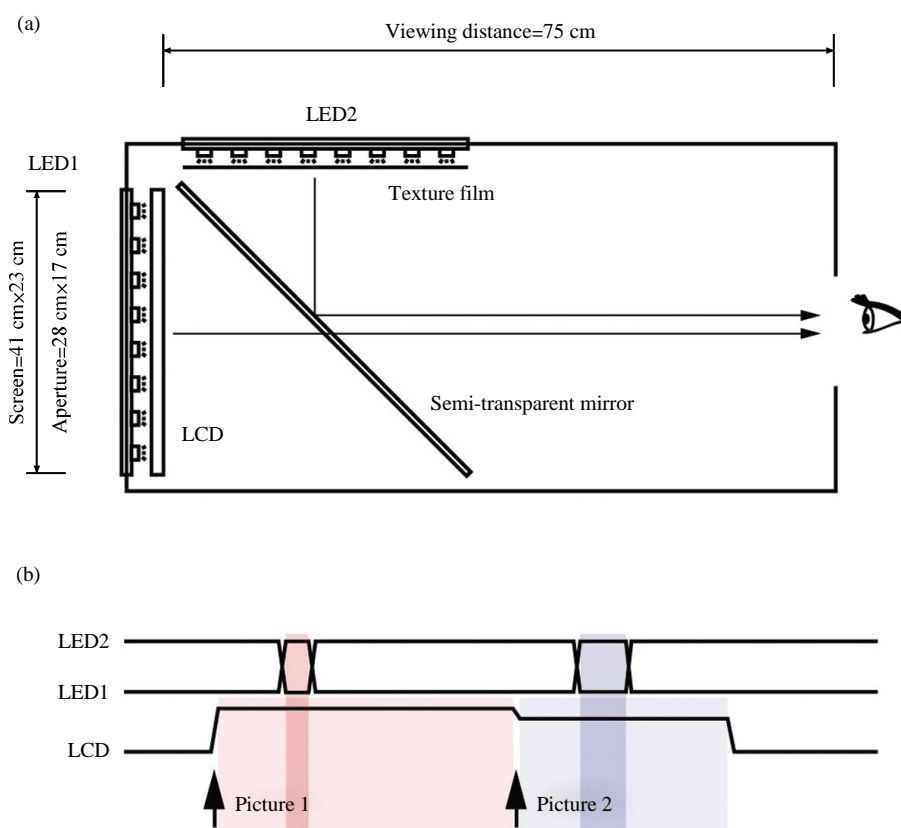


Fig. 1 LED backlight tachistoscope setup and timing diagram

(a) The tachistoscope comprises screen1 (LCD and LED1) and screen2 (texture film and LED2), separated by a semi-transparent mirror in a light-proof box. The relative position of the three components are adjusted to create the impression of a single screen 75 cm away when observed from the viewing window. (b) After a new picture is updated on the LCD panel (light shaded), it is not visible unless LED1 is turned on. Stimulus exposure duration (dark shaded) is determined by the short LED1 pulse, not affected by the sluggish and asynchronous nature of LCD. LED2 is in the anti-phase of LED1, so that the texture on the projector film will be visible when LED1 is turned off.

1.3 时序控制程序

为保证呈现时间亚毫秒级的精确性，LED 背光的翻转时间完全由单片机控制。每个试次开始前，计算机通过串口指令向单片机发送本次需要的呈现时间。当要呈现的图片翻转到显卡的前端缓存，开始向显示器传输时，计算机向单片机发送开始信号。我们测试了在开始信号到达后不同时间点打开 LED1 时图像的可见性，发现大约在 50 ms 后（约 3 个刷新周期）图像才达到稳定。保守起见，正式实验中我们在开始信号到达 100 ms 后再打开 LED1。之后，单片机延迟事先指定的时间后自动关闭 LED1，无需计算机控制，确保实验程序及通信延迟不会影响呈现时间的精确性。单片机的计时精度达到微秒级，LED 的开关响应时间也在亚毫

秒级，所以我们的速视器能够实现小于 1 ms 的短暂呈现，呈现时间也能以小于 1 ms 的时间精度连续调节，而且整幅图像可以做到同时呈现，同时消失。

1.4 亮度匹配与掩蔽效应

采用最小闪烁法(minimum flicker procedure)匹配 2 块屏幕的亮度，可以实现无掩蔽的快速呈现。然而在本实验中，我们希望使用后向掩蔽技术更好地控制刺激的加工时间。模式掩蔽(pattern masking)和自身对比掩蔽(metacontrast masking)在时间过程上通常表现为 U 形曲线，涉及较为复杂的认知机制^[28]。因此这里我们将 LED2 的亮度调至最大值，使用简单的高亮度白屏作为掩蔽刺激。

2 图形结构优势效应

2.1 材料与方法

2.1.1 受试.

12 名受试(年龄 20~29 岁, 其中 4 位女性, 1 人是本文作者, 均为右利手)自愿参加了实验. 所有受试均报告视力正常或矫正后正常, 并在实验开始前给予了知情同意. 除本文作者外, 所有受试事前均不了解实验目的, 并在实验结束后接受了一定数量的受试费作为交通和时间补偿.

2.1.2 刺激.

刺激图形如图 2 所示, 为白底黑色线条图形, 整体占据约 $9^\circ \times 9^\circ$ 视角范围. 刺激分布在中央注视点周围的 4 个象限中, 每个象限包含 4 个子图形. 实验包含欧氏、仿射、射影和拓扑 4 个条件. 每个条件下, 其中 3 个象限的图形具有相同的该种几何不变性, 剩下的目标象限与其他象限不同. 对于欧氏条件, 目标象限箭头的朝向角度与其他象限不同; 对于仿射条件, 目标象限线段的平行关系与其他象限不同; 对于射影条件, 目标象限的共线性关系(折线段还是直线段)与其他象限不同; 对于拓扑条件, 目标象限的连通性以及洞的个数与其他象限不同(4 个直角元素可能联通为一个方块). 每种条件下, 目标象限都可以是 1~4 象限中的任何一个. 此外, 每种不变性的两种状态都可以出现在目标象限中. 例如对于拓扑条件, 既可能是 3 个折线

1 个方块, 也可能是 3 个方块 1 个折线. 两种状态出现的概率相等, 受试无法只依据特定的形状做出判断, 而必须比较 4 个象限的异同. 每种条件共包含 8 幅不同的刺激图形, 在实验中等概率随机出现.

2.1.3 流程.

每个受试都需要完成两个阶段的实验. 阶段 1 和阶段 2 的刺激图形和试次结构完全相同, 唯一不同的是决定每个试次刺激呈现时间的方法. 受试被要求首先盯住屏幕白色背景中央的黑色注视点, 准备好后按下键盘空格键自己启动新的试次. 按键后, 经过 1000~2000 ms 的随机间隔, 屏幕上会快速闪现一幅刺激图形. 刺激消失后, 屏幕上继续显示注视点. 由于注视点的白色背景由亮度更高的 LED2 显示, 它也起到了前向和后向掩蔽刺激的作用. 受试的任务是在刺激图形消失后, 尽可能准确地按键报告是哪一个象限与其他象限图形性质不同. 为了防止不小心按错键造成的失误, 受试在按空格键启动下一个试次之前, 可以重新按键修改自己的选择.

在正式实验开始前, 受试会在几个不同的呈现时间上进行少量的练习, 直至他们对实验任务和反应模式感到适应.

在阶段 1, 我们采用混合阶梯法改变每个试次的刺激呈现时间, 每种条件的呈现时间由一个独立的阶梯控制. 阶梯采用 2 下 1 上规则, 每次调整将呈现时间缩短或延长 1dB(1.122 倍). 4 种条件等概率随机出现, 直到所有条件都完成 40 个试次. 阶段 1 结束后, 我们会估计每个受试在 4 种条件下的呈现时间阈值, 并据此确定该受试阶段 2 中呈现时间的范围.

在阶段 2, 我们采用混合条件的固定刺激法, 在阶段 1 确定的呈现时间范围内按对数取等间距的 6 个不同呈现时间. 每种条件的每个呈现时间各重复 16 次, 以等概率随机出现. 总共 384 个试次, 分成两大组完成. 第一大组结束后, 我们会初步绘制受试的心理物理曲线, 如果发现呈现时间范围选择不理想, 进行第二大组时还会稍加调整.

将实验分成两个阶段进行是为了最好地根据每个受试的情况自适应地选择固定刺激法的测量范围. 将 4 种条件逐试次地混在一起随机呈现是为了防止受试事先将注意集中在某种图形特征上, 同时也可以较好地在条件间平衡练习与疲劳效应.

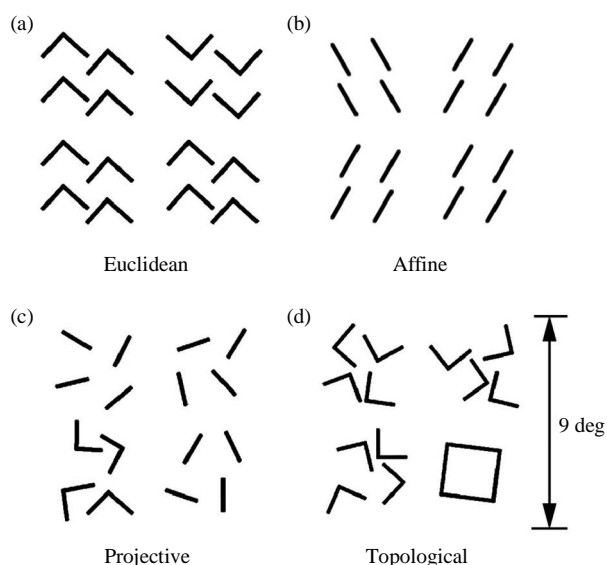


Fig. 2 Examples of visual display in each condition

The odd quadrant in the example display is the 1st~4th quadrant for (a)~(d), respectively. The fixation point was in the center of the display (not shown).

2.1.4 数据分析与统计.

我们主要分析阶段 2 得到的数据. 将实际呈现时间取自然对数, 采用最大似然估计拟合逻辑斯蒂函数(logistic function)形式的心理物理曲线, 假定随机水平为 25%, 失误率为 1%, 每个受试的每种条件分别拟合, 拟合结果的置信区间采用自助法(bootstrap)随机重采样 1 000 次进行估计. 阈值定义为正确率为 70.7%时的呈现时间, 然后对每种条件的阈值进行随机效应的组分析.

组分析时首先对 4 种不同的几何不变性进行单因素重复测量的方差分析. 主效应显著时, 再进行两两配对 t 检验, 检验结果的 P 值采用 Holm 多重比较校正.

2.2 结果与讨论

我们首先考察了目标象限位置对整体正确率是否存在影响. 合并其他因素后, 单因素重复测量的

方差分析结果表明象限的主效应不显著, $F(3, 33)=0.503$, $P=0.683$. 由于在本实验中我们并不关心刺激位于视野不同象限对不同几何性质加工速度的影响, 同时也缺少足够的试次数来分离象限与其他因素的交互作用, 因此在后续分析中合并了 4 个象限的数据.

图 3 显示了一个典型受试的心理物理曲线. 对于 4 种几何不变性, 受试对目标象限判断的正确率都受到呈现时间的显著调制, 且可由 logistic 函数较好地刻画. 这表明通过限制刺激的加工时间, 可以用心理物理曲线较好地描述对几何不变性差异的检测过程. 注意到正确率随呈现时间变化得非常快, 大约在 15 ms 之内就从随机水平上升到了接近饱和. 这种快速的动态过程充分体现了可精细调整呈现时间的新型速视器对本实验的重要性.

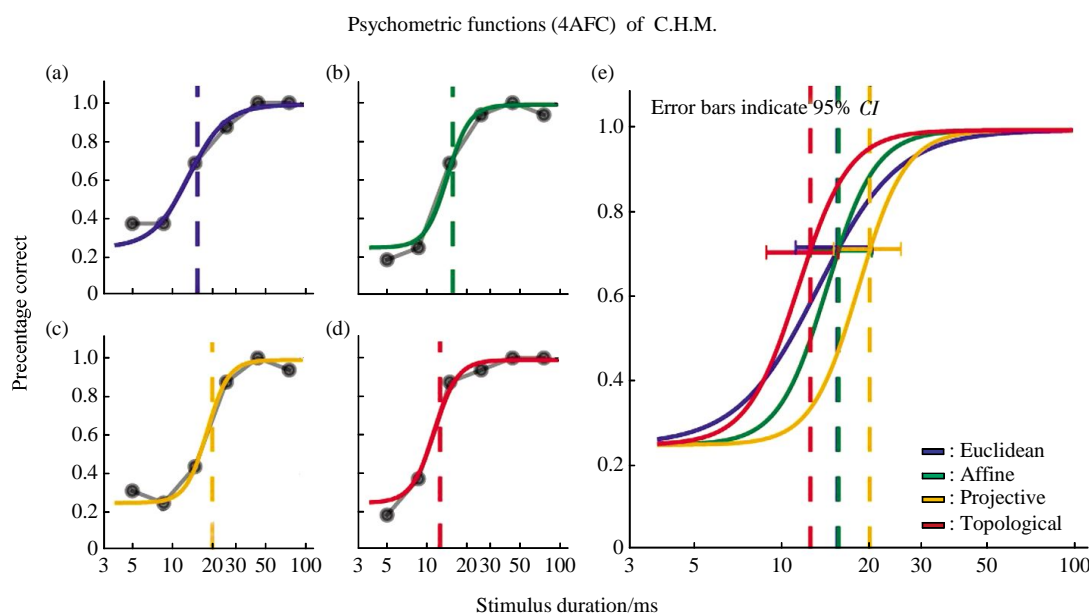


Fig. 3 Psychometric function in each condition for one subject

(a) Euclidean. (b) Affine. (c) Projective. (d) Topological. (e) The same four psychometric functions plotted together for easier comparison. The dashed vertical lines indicate stimulus duration threshold at 70.7% percentage correct. The error bars indicate 95% confidence interval of the threshold estimated using bootstrap procedure.

加工不同几何不变性需要的知觉整合时间明显不同, 如图 4a 所示. 单因素重复测量的方差分析表明不变性的主效应显著, $F(3, 33)=11.43$, $P<0.001$. 进一步的两两配对 t 检验表明, 除欧氏性质与射影性质的差异不显著外, 其余性质的呈现时间阈值差

异均达到统计显著或临界显著. 欧氏性质需要的整合时间((24±11) ms)大于仿射性质((15±6) ms), 仿射性质大于拓扑性质((12±4) ms), 符合大范围首先理论对功能层次的预言. 拓扑性质需要的整合时间最短, 大约只有欧氏性质的 1/2.

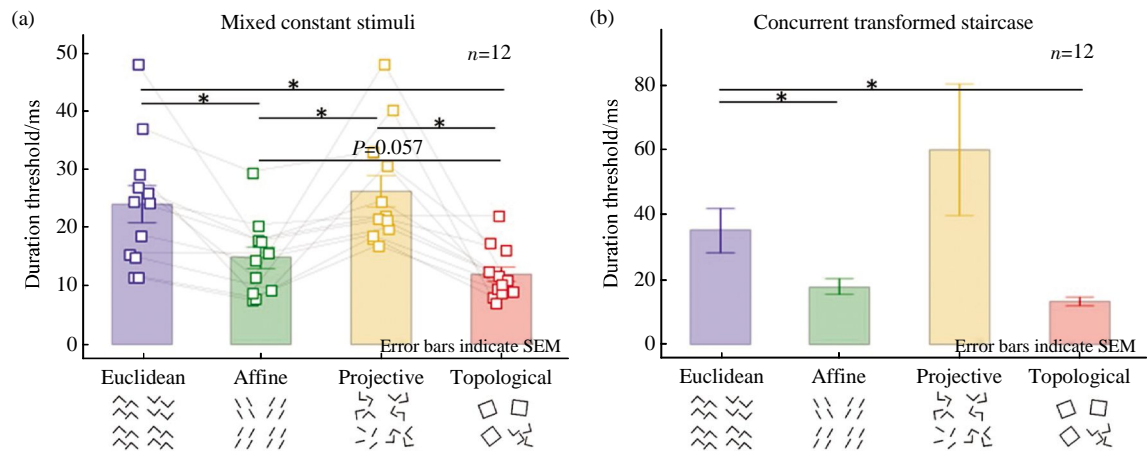


Fig. 4 Group analysis results for stage 2 (mixed constant stimuli) and stage 1 (concurrent transformed staircase)

(a) Stimulus duration threshold for each condition in stage 2 (main results). Each marker represents the threshold data from one subject. The light gray lines connect data points from the same subject across conditions. (b) Stimulus duration threshold for each condition in stage 1 (pilot data). For both panels, the error bars indicate standard error of the mean.

我们也用类似的方法分析了阶段 1 的数据, 如图 4b 所示。虽然由于试次数较少使得数据波动较大, 但结果的趋势与阶段 2 完全一致, 这进一步验证了上述结果的可靠性。单因素重复测量的方差分析表明不变性的主效应显著, $F(3, 33)=4.035$, $P=0.015$ 。进一步的两两配对 t 检验表明, 欧氏性质与拓扑性质及仿射性质的差异统计显著。

值得注意的是, 在上述结果中射影性质所需的整合时间最长, 与欧氏性质在统计上无显著差异, 这与功能层次的预言有一定分歧。我们将在本文第 3 部分讨论可能的原因。

3 讨 论

3.1 拓扑性质的优先加工

动物要在自身和环境的不断运动与变化中保持稳定的知觉, 就需要利用在变换下保持不变的性质。稳定性越高的性质, 对于表征物体具有越基础性的作用。长度和角度这样的欧氏性质在小幅度的平面内平移和旋转中保持不变, 但当与被观察物体的距离(视距)发生变化时, 看到的长度就会发生变化; 平行性这样的仿射性质不受视距影响, 但当观察的视角发生显著变化时, 原先平行的轮廓看起来就不再平行; 共线性这样的射影性质在不同视角下都保持不变, 但假如物体能够改变形状(例如小鸟扇动翅膀), 射影性质也会被破坏。稳定性最高的

是连通性、连通支集的个数以及洞的个数这样的拓扑性质。按照大范围首先理论, 图形的拓扑性质会在视觉过程的早期通过某些快速加工通道被优先抽提出来, 用以建立知觉物体的表征。我们对不同几何不变性加工所需最短呈现时间的实验结果表明, 拓扑性质分辨需要的呈现时间最短, 处理速度最快, 这与大范围首先理论的预测是一致的, 与早先的反应时研究^[8]互为补充。

3.2 变换稳定性与结构稳定性

在刺激呈现时间不断缩短最后达到仅有几毫秒的过程中, 受试的主观感觉是刺激图形似乎变得碎片化地局部不可见, 有时只能随机看到刺激图形的一部分, 类似于刺激信号被随机图像噪声污染的效果。事实上, 造成这种现象的可能原因之一就是短暂呈现使得感觉输入信号十分微弱, 部分淹没在感觉编码神经系统的生理噪声当中。这种对图形的劣化是不同于 Erlangen 纲领中几何变换的另一种变换。按照近年来兴起的拓扑数据处理理论, 我们可以用结构稳定性的概念来描述, 图形在局部被不同尺度的噪声破坏的情况下, 是否能保持原有的性质。从进化的角度看, 优先处理在噪声干扰下仍能保持的信息对生物体具有重要的意义, 因此变换稳定性和结构稳定性可能在功能层次上具有某种关联。

在本实验中, 完成 CSE 任务的正确率与图形

在被噪声劣化的情况下能否保持相关的几何不变性有关。目前的结果似乎表明,拓扑性质具有较高的结构稳定性。而射影性质的阈值较高,提示两种稳定性的层次关系也许不完全一致。这些推测还有待使用类型更丰富多变的刺激图形来进一步检验。

3.3 布洛赫定律(Bloch's law)与时间整合

在极短的呈现时间条件下,受试的另一种主观感觉是刺激图形的可分辨性不断下降,表现之一是亮度和对比度的降低(在没有掩蔽刺激的条件)。那么我们是否可以通过调整刺激的亮度和对比度来达到同样的效果,从而在普通显示器上完成类似的实验呢?

Graham 和 Margaria^[29]发现对于视角很小的刺激,当呈现时间短于 10 ms 时,检测阈值附近的刺激亮度和呈现时间呈现倒数关系。换言之,是亮度和时间的乘积决定了刺激是否能够通过阈值,这个结果被称为布洛赫定律^[30]。看起来,我们似乎可以通过降低亮度在更长的呈现时间下取得类似的效果。

然而,上述结果最初描述的主要是视网膜光感受器的时间整合特性。而在复杂刺激图形的辨别任务中,时间整合也发生在决策系统对知觉线索的累积过程中。此时知觉线索的强度和呈现时间可能仍然具有某种可置换性,但其机制却不一样。Greene 及其同事^[30]发现,对于形状辨识任务,亚毫秒级呈现时布洛赫定律并不成立。这表明缩短呈现时间可能不能简单地由降低亮度或者对比度来代替,而高速显示设备在视知觉研究中具有独特的作用。

3.4 视觉短时记忆

虽然在实验中采用高亮度的白屏作为掩蔽刺激,试图削弱视觉短时记忆对任务的影响,但我们仍然不能完全排除这一点。我们采用的任务本质上是一个定位目标象限的知觉决策任务,类似于视觉搜索。如果定位不同类型的不变性差异所需要的搜索反应时不同,那么它们受到视觉短时记忆消退过程的影响就不同,进而影响不同条件下的正确率。

当然,不变性差异对搜索效率的影响本身就是一个有趣的问题,也给出了对不同几何不变性加工机制的重要洞见。大范围首先理论对视觉搜索结果的预测与本文结果并不矛盾。为了进一步减少与我们关心的认知过程无关的复杂知觉决策过程带来的干扰,将来的实验可以采用更为简单的任务,例如要求受试无需指出是哪个象限不同,只需判断是否存在一个象限与其他象限不同。

4 结 论

我们自行研制了新型 LED 背光速视器,并利用其允许精细控制视觉刺激短暂呈现的特点,采用限制刺激加工时间的心理物理方法,研究了 CSE 任务下加工不同几何不变性所需要的知觉整合时间。结果表明,拓扑性质的处理速度最快。同时,心理物理方法克服了反应时实验的一些缺点,它与新型速视器的结合应用,可能为认知过程动态特性的研究提供新的有效手段。

致谢 感谢李志光、左真涛和张紫豪在液晶显示器改造过程中提供的宝贵帮助和建议。感谢胡鲲在实验过程中对我们研究的支持。

参 考 文 献

- [1] Cox D D. Do we understand high-level vision?. *Current opinion in neurobiology*, Elsevier Ltd, 2014, **25C**: 187-193
- [2] Rajalingham R, Schmidt K, DiCarlo J J. Comparison of object recognition behavior in human and monkey. *Journal of Neuroscience*, 2015, **35**(35): 12127-12136
- [3] DiCarlo J J, Cox D D. Untangling invariant object recognition. *Trends in Cognitive Sciences*, 2007, **11**(8): 333-341
- [4] DiCarlo J J, Zoccolan D, Rust N C. How does the brain solve visual object recognition?. *Neuron*, Elsevier Inc., 2012, **73**(3): 415-434
- [5] Kravitz D J, Saleem K S, Baker C I, *et al.* The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, Elsevier Ltd, 2013, **17**(1): 26-49
- [6] Serre T, Oliva A, Poggio T A. A feedforward architecture accounts for rapid categorization. *Proc Nat Acad Sci USA*, 2007, **104**(15): 6424-6429
- [7] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks. *NIPS*, 2012: 1-9
- [8] Yamins D L K, Hong H, Cadieu C F, *et al.* Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc Nat Acad Sci USA*, 2014, **111**(23): 8619-8624
- [9] Hochstein S, Ahissar M. View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron*, 2002, **36**(5): 791-804
- [10] Chen L. The topological approach to perceptual organization. *Vis. Cogn.*, HOVE: PSYCHOLOGY PRESS, 2005, **12**(4): 553-637
- [11] Pomerantz J R, Sager L C, Stoeve R J. Perception of wholes and of their component parts: some configural superiority effects. *Journal of experimental psychology. Human Perception and Performance*, 1977, **3**(3): 422-435
- [12] Pomerantz J R, Portillo M C. Grouping and emergent features in vision: toward a theory of basic Gestalts. *Journal of Experimental Psychology: Human Perception and Performance*, 2011, **37** (5): 1331-1349

- [13] Pomerantz J R. Are complex visual features derived from simple ones? Leeuwenberg E L J, Buffart H F J M. Formal theories of visual perception. New York: John Wiley & Sons, 1978: 217–229
- [14] Kveraga K, Boshyan J, Bar M. Magnocellular projections as the trigger of top-down facilitation in recognition. *J Neurosci*, 2007, **27**(48): 13232–13240
- [15] Zhou T, Zhang J, Chen L. Neural correlation of "global-first" topological perception: anterior temporal lobe. *Brain Imaging and Behavior*, 2008, **2**(4): 309–317
- [16] Lu Z-L, Doshier B. Visual Psychophysics: From Laboratory to Theory. USA: MIT Press, 2014
- [17] Pachella R G. The interpretation of reaction time in information processing research//Kantowitz B H. Human Information Processing: Tutorials in Performance and Cognition. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974: 41–82
- [18] Wang B, Zhou T G, Zhuo Y, *et al.* Global topological dominance in the left hemisphere. *Proc Nat Acad Sci USA*, 2007, **104** (52): 21014–21019
- [19] Orsten-Hooge K D, Portillo M C, Pomerantz J R. False pop out. *J Exp Psychol. Hum Percept Perform*, 2015, **41**(6): 1623–1633
- [20] Luce R. Response times?: their role in inferring elementary mental organization. New York; Oxford: Oxford University Press; Clarendon Press, 1986
- [21] Wickelgren W A. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 1977, **41**(1): 67–85
- [22] Stanford T R, Shankar S, Massoglia D P, *et al.* Perceptual decision making in less than 30 milliseconds. *Nature neuroscience*, Nature Publishing Group, 2010, **13**(3): 379–385
- [23] Bacon-Macé N, Macé M J M, Fabre-Thorpe M, *et al.* The time course of visual processing: Backward masking and natural scene categorisation. *Vision Research*, 2005, **45**(11): 1459–1469
- [24] Thurgood C, Patterson J, Simpson D, *et al.* Development of a light-emitting diode tachistoscope. *Review of Scientific Instruments*, 2010, **81**(3): 035117
- [25] Thurgood C, Whitfield T W A, Patterson J. Towards a visual recognition threshold: new instrument shows humans identify animals with only 1ms of visual exposure. *Vision Research*, Copyright (c) 2011 Elsevier Ltd. All rights reserved, 2011, **51**(17): 1966–1971
- [26] Sperdin H F, Repnow M, Herzog M H, *et al.* An LCD tachistoscope with submillisecond precision. *Behavior research methods*, 2013, **45**(4): 1347–1357
- [27] Sperdin H F, Spierer L, Becker R, *et al.* Submillisecond unmasked subliminal visual stimuli evoke electrical brain responses. *Human Brain Mapping*, 2015, **36**(4): 1470–1483
- [28] Breitmeyer B, Ögmen H. Visual masking: time slices through conscious and unconscious vision. 2nd. Oxford; New York: Oxford University Press, 2006
- [29] Graham C H, Margaria R. Area and the intensity-time relation in the peripheral retina. *American Journal of Physiology, Am Physiological Soc*, 1935, **113**(2): 299–305
- [30] Greene E, Ogden R T. Shapes displayed with durations in the microsecond range do not obey Bloch's law of temporal summation. *i-Perception*, 2013, **4**(6): 429–436

High Speed LED Backlight Tachistoscope Reveals Fast Processing of Topological Invariants*

QIAN Chen-Can^{1,2)}, LIU Zu-Xiang¹⁾**

⁽¹⁾ Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China;

⁽²⁾ University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract Primate visual system is extremely efficient in extracting stable invariants from environment, despite various transformation and degradation. The timing and dynamics of invariants perception provide important constraints for theories of vision. We developed a high speed LED backlight tachistoscope capable to deliver visual stimuli with 1 ms exposure and adjust the duration in submillisecond precision, which enabled us to investigate the relative speed of processing topological, projective, affine, and Euclidean invariants in configural superiority effect paradigm by manipulating the access time to visual input and estimating the psychometric function for sensory evidence accumulation. The results suggest that topological invariant requires shortest exposure time, consistent with the prediction of global-first theory.

Key words configural superiority effect, topological invariant, timing, psychometric function, tachistoscope
DOI: 10.16476/j.pibb.2015.0300

*This work was supported by grants from National Basic Research Program of China (973 grant)(2012CB825500, 2015CB351701), the National Natural Science Foundation of China (91132302), and Science Foundation of the Chinese Academy of Sciences grants (XDB02010001, XDB02050001).

**Corresponding author.

Tel: 86-10-64888565-808, E-mail: zxliu@bcslab.ibp.ac.cn

Received: September 29, 2015 Accepted: January 7, 2016